

Article

## The *Apis mellifera* Filamentous Virus Genome

Laurent Gauthier <sup>1,†,\*</sup>, Scott Cornman <sup>2,†,‡</sup>, Ulrike Hartmann <sup>1</sup>, François Cousserans <sup>1</sup>, Jay D. Evans <sup>2</sup>, Joachim R. de Miranda <sup>3</sup> and Peter Neumann <sup>1,4</sup>

<sup>1</sup> Agroscope, Swiss Bee Research Centre, Schwarzenburgstrasse 161, CH-3003 Bern, Switzerland; E-Mails: uli.hartmann@gmx.ch (U.H.); coussera@supagro.inra.fr (F.C.); peter.neumann@vetsuisse.unibe.ch (P.N.)

<sup>2</sup> Bee Research Laboratory, Beltsville, MD 20705, USA; E-Mails: scott.cornman@gmail.com (S.C.); jay.evans@ars.usda.gov (J.D.E.)

<sup>3</sup> Department of Ecology, Swedish University of Agricultural Sciences, Uppsala 750 07, Sweden; E-Mail: joachim.de.miranda@slu.se

<sup>4</sup> Institute of Bee Health, Vetsuisse Faculty, University of Bern, CH-3001 Bern, Switzerland

† These authors contributed equally to this work.

‡ Current address: Fort Collins Science Center, USGS, Fort Collins, CO 80526, USA.

\* Author to whom correspondence should be addressed; E-Mail: laurent.gauthier@agroscope.admin.ch; Tel.: +41-584-622-155; Fax: +41-584-638-227.

Academic Editors: Elke Genersch and Sebastian Gisder

Received: 31 March 2015 / Accepted: 2 July 2015 / Published: 9 July 2015

**Abstract:** A complete reference genome of the *Apis mellifera* Filamentous virus (AmFV) was determined using Illumina Hiseq sequencing. The AmFV genome is a double stranded DNA molecule of approximately 498,500 nucleotides with a GC content of 50.8%. It encompasses 247 non-overlapping open reading frames (ORFs), equally distributed on both strands, which cover 65% of the genome. While most of the ORFs lacked threshold sequence alignments to reference protein databases, twenty-eight were found to display significant homologies with proteins present in other large double stranded DNA viruses. Remarkably, 13 ORFs had strong similarity with typical baculovirus domains such as PIFs (*per os* infectivity factor genes: *pif-1*, *pif-2*, *pif-3* and *p74*) and BRO (Baculovirus Repeated Open Reading Frame). The putative AmFV DNA polymerase is of type B, but is only distantly related to those of the baculoviruses. The ORFs encoding proteins involved in nucleotide metabolism had the highest percent identity to viral proteins in GenBank. Other notable

features include the presence of several collagen-like, chitin-binding, kinesin and pacifastin domains. Due to the large size of the AmFV genome and the inconsistent affiliation with other large double stranded DNA virus families infecting invertebrates, AmFV may belong to a new virus family.

**Keywords:** *Apis mellifera*; filamentous virus; genome annotation

---

## 1. Introduction

*Apis mellifera* filamentous virus (AmFV) was originally described as a rickettsia disease of honeybees (*Apis mellifera*) [1], but was later characterized as a large enveloped DNA virus [2]. It derives its name from the long ( $3150 \times 40$  nm) filamentous nucleoprotein that folds in three superimposed figure-8 loops (Figure 1) into a  $450 \text{ nm} \times 170 \text{ nm}$  rod-shaped virion [2,3]. The nucleoprotein contains a central core of double stranded DNA (dsDNA) wrapped by two major nucleoproteins while the tri-laminate virion envelope contains lipids, two major proteins and several minor proteins [4]. One major diagnostic feature of AmFV infection is that the hemolymph of severely infected adult honeybees becomes milky-white due to cellular degradation and the large number of virions present, which is also a primary feature of insect ascoviruses infecting Lepidopteran hosts [5].

*A. mellifera* filamentous virus epizootics were identified in various parts of the world through electron microscopic examination of bee hemolymph samples [6]. AmFV was shown to multiply in different bee tissues, principally in the fat body and ovarian tissues of adult bees [7]. Highly infected bees display signs of weakness and are usually gathered at the hive entrance. Although it may occasionally induce such overt, colony-level symptoms, AmFV is usually considered a weakly pathogenic virus of honeybees, with low impact on host lifespan [7,8]. Experimental infections can be established by feeding adult bees AmFV particles, but only when co-infected with *Nosema apis* spores [9]. Historically, there has been a strong association between *Nosema sp.* and AmFV infections [7]. AmFV has occasionally been associated with winter honeybee colony mortality [3], but its annual peak incidence is mostly in spring, coinciding with the peak incidence of *N. apis* [9]. Filamentous viruses with similar morphology and tissue distribution have also been observed in a number of parasitic wasps, where they may be associated with super-parasitism behavior [10].

Here we present the complete nucleotide sequence and annotation of the AmFV genome, as well as insights into the distribution of this virus in honeybee colonies. Although some coding sequences are homologous with other large dsDNA viruses of invertebrates, such as baculoviruses, the size of AmFV genome, the proportion of putatively non-coding sequences and the large number of open reading frames (ORFs) lacking database matches all suggest a very distinct virus that merits classification in a new family of insect viruses.

## 2. Materials and Methods

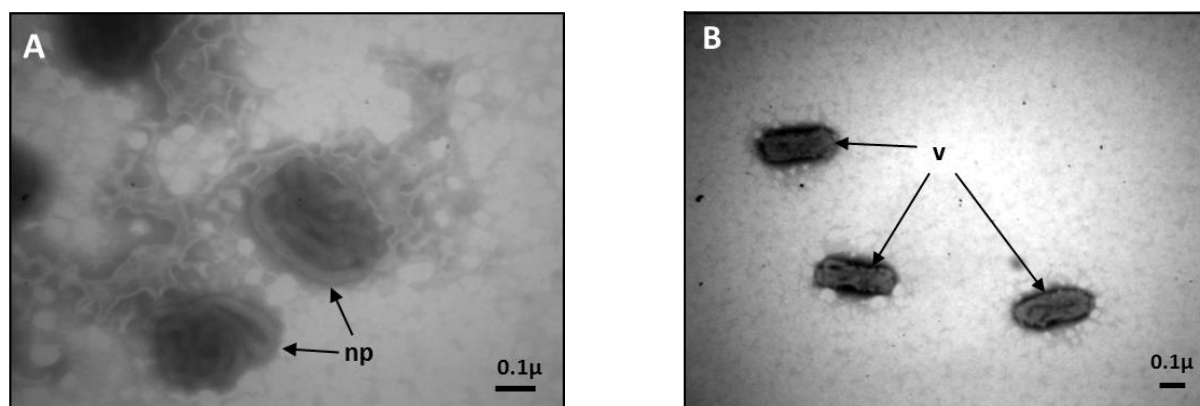
### 2.1. AmFV DNA Isolation, Sequencing, and Assembly

Viral particles were purified from six infected honeybee workers, collected from a single colony in the spring of 2009 in Bern, Switzerland, that were displaying the characteristic “milky hemolymph”

AmFV clinical signs. After homogenization in 1.5 mL of 50 mM Tris pH 7.5 buffer, viral particles were separated by centrifugation on a 10%–60% Nicodenz<sup>®</sup> (Sigma-Aldrich, St Louis, MI, USA) gradient. Two visible bands were extracted from the gradient and analyzed by transmission electron microscopy (TEM). The presence of AmFV enveloped virions (Figure 1) was confirmed in the lower band while the upper band contained mostly degraded nucleoproteins. Total DNA was purified from the lower band by phenol-chloroform extraction and precipitation with ethanol. The DNA was resuspended in 5 mM Tris pH 7.5 and quantified using the Qubit<sup>®</sup> system (Life Technologies, Carlsbad, CA, USA). One hundred nanograms of DNA were sent to Fasteris Life Science Co. (Geneva, Switzerland) where a paired-end library was prepared using the TruSeq<sup>®</sup> Nano DNA LT kit (Illumina, San Diego, CA, USA) with bead-based size selection of 550 bp fragments. The library was sequenced on an Illumina HiSeq 2500 Analyzer, producing 100-bp paired reads that were quality filtered and adapter screened by the sequence provider and assembled *de novo* at a range of kmers using the VELVET program [11].

Alignments from the scaffold dataset were produced using the SeqMan (Lasergene 7.0; Dnastar Inc., Madison, WI, USA) software. Gaps and ambiguous regions were amplified by PCR using a high fidelity polymerase (HIFI, Kapa-Biosystems, Wilmington, MA, USA) and appropriate primers. The resulting amplicons were sequenced on both strands using the Sanger sequencing method.

Jellyfish [12] was used to evaluate the heterogeneity of the isolated sample by plotting the coverage of all kmers of length 59 occurring in the raw reads. Batches of kmers were assembled with CLC Genomics v. 7.1 (Qiagen, Valencia, CA, USA) using a kmer of 59 and searched by BLASTX against the UniRef90 database [13].



**Figure 1.** Morphology of AmFV virions: (A,B) Electron micrographs of “milky” bee hemolymph containing characteristic AmFV nucleoproteins (np) and enveloped virions (v). Approximate size is indicated by scale bars.

## 2.2. AmFV Sequence Analysis

Methionine-initiated open reading frames (ORFs) of 100 amino acids or more were identified with the getorf program of the EMBOSS package [14]. Those ORFs with no or minimal overlap were identified using the AMIGENE gene prediction software [15] and were assumed to encode putative proteins. The sequences of these ORFs were compared with BLASTP to the “UniRef90” database and with delta-blast to the conserved domain database (CDD) of NCBI, in both cases with a minimum e-value of  $1.0e^{-5}$ , as well as against the Pfam-A database (Hmmer3) with a minimum e-value for the best domain match

of 0.1. The graphical representation of genomic ORFs was performed using GenQuest (Lasergene 7.0; Dnastar Inc., Madison, WI, USA). Multiple sequence alignments were performed using the Multalin software [16]. Phylogenetic analyses were conducted using the pipeline offered by Phylogeny.fr: MUSCLE was used for multiple alignment, Gblocks for alignment curation, PhyML for phylogeny and finally TreeDyn for tree drawing [17]. Pfam domains of proteins were identified using SMART [18]. The GC content of the genome was calculated using DNA plotter [19].

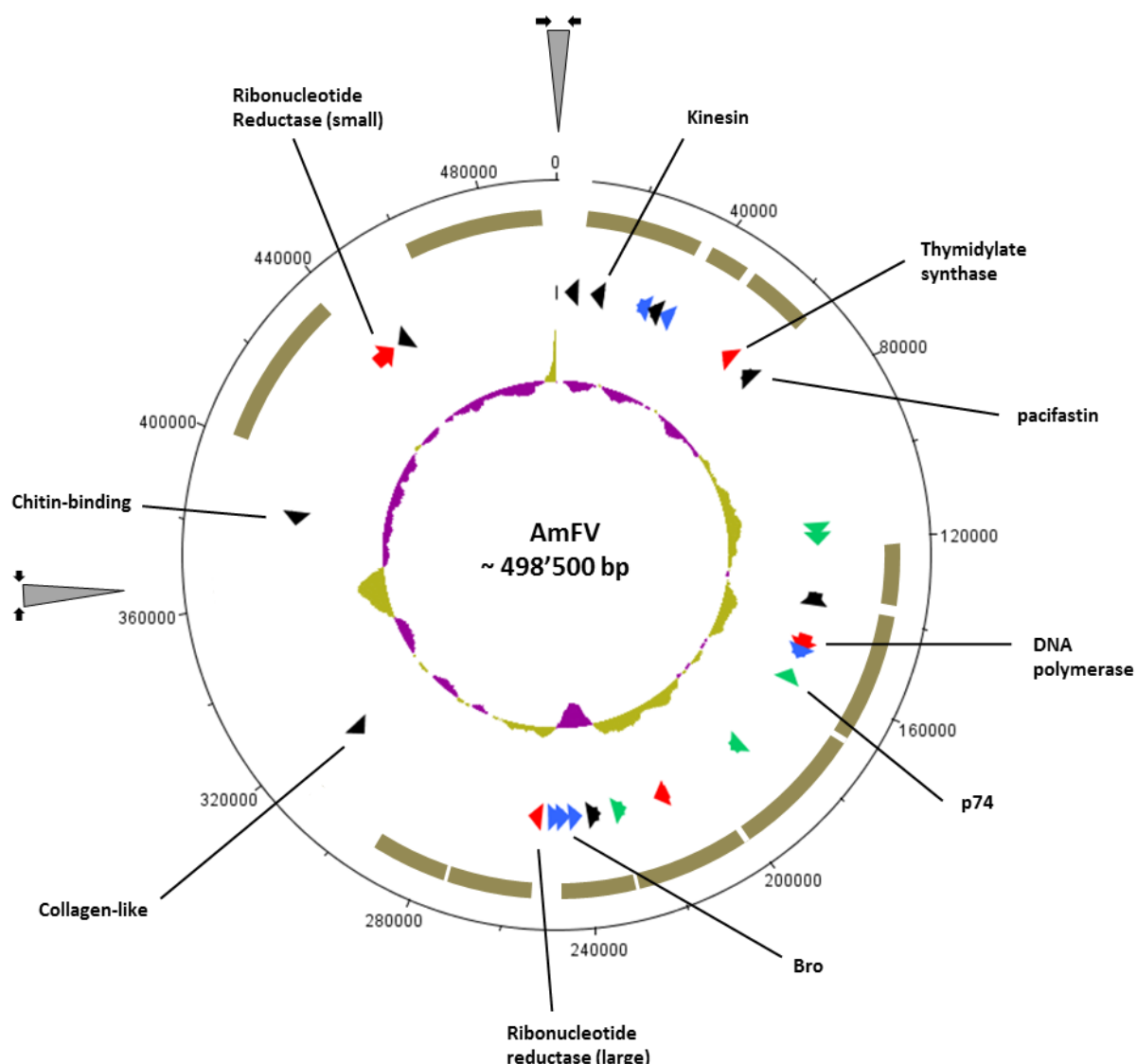
### 2.3. Assembly of AmFV-Like Contigs from the USA

Cornman *et al.*, 2010 [20] identified several putative virus sequences from shotgun genomic sequencing of a U.S.A. sample of the bee parasite *Varroa destructor*, which, by BLAST alignment, were found to partly overlap with the AmFV genome sequence presented here. The original reads (N = 95,386) underlying those varroa-derived contigs were reassembled with Mira4 [21], in order to identify and maximize the major contigs for this virus isolate, resulting in 844 contigs of 300 bp or more, with an N50 of 2052 bp. MegaBLAST was used to select a subset of contigs matching the AmFV reference assembly presented here. These 88 contigs totaled 461.7 kb of sequence with an average G + C content of 50.0%. The coverage and consistency of the final genome model was assessed by mapping read pairs with “Bowtie2” using the “very-sensitive” and “end-to-end” quick switches. The statistics for this alignment were calculated with SAMtools [22] and the Picard package [23].

### 2.4. AmFV PCR Detection

Adult and larval honeybees were collected from their colony and analyzed either individually or in pools of 100 workers. Organs were obtained after dissection in PBS under a stereomicroscope, as described previously [24]. Bee products (honey and pollen) were directly collected from the frames. Bees were homogenized in Tris 20 mM, pH 7.5 – NaCl 400 mM buffer using a 5 mm metal bead and the Tissue Lyzer<sup>®</sup> homogenizer (Qiagen, Valencia, CA, USA) for 30 s, while bee products and organs were directly homogenized following the same process in a denaturing buffer (Buffer T1, Nucleospin Tissue<sup>®</sup> kit, Macherey-Nagel, Düren, Germany). Fifty microliters of the homogenate was used for DNA extraction, using the NucleoSpin<sup>®</sup> Tissue kit (Macherey Nagel, Düren, Germany) following the manufacturer’s recommendations. The DNA was eluted in 100 µL of water and quantified by spectrophotometry (Nanodrop<sup>®</sup>, Wilmington, DE, USA) prior to being used for PCR analysis. The quality of the DNA samples was confirmed by quantifying the *A. mellifera* actin gene in the samples using quantitative PCR as described previously [25]. Diagnostic PCR primers were designed for 3 genes displaying 100% nucleotide identity between the Swiss and US AmFV sequences and located in different regions of the AmFV genome (see Figure 2 and Suppl. File 1): a thymidylate synthase gene (AmFV\_28: forward CGCATGTACCAACAACCTCGTAC and reverse CACAGTTGGTGTAGCGCAGT, described by Cornman *et al.*, 2010 [20]), a ribonucleotide reductase gene (AmFV\_221: forward AGATCGCCCGCTTTGTGCGC and reverse AGCGGGCCTCGGTGTACACT) and a Bro gene (AmFV\_112: forward CAGAGAATTCGGTTTTTGTGAGTG and reverse CATGGTGGCCAAGTCTTGCT). PCR reactions were performed using standard reagents containing 0.4 µM of each primer and a cycling profile of 2 min at 94 °C followed by 35 cycles of: 30 s at 94 °C,

30 s at 56 °C and 45 s at 72 °C. The identity of the PCR amplicons was confirmed using Sanger sequencing.



**Figure 2.** Schematic representation of the AmFV genome. Characteristic ORFs presenting similarities to other large dsDNA viruses are indicated. Colors refer to putative functions: DNA replication and nucleotides metabolism genes (red), *pif* and *p74* homologs (green), *Bro* genes (blue), and putative virulence genes (black). The brown circle fragments correspond to the positions of the larger contigs (>8500 bp) filtered out from *V. destructor* metagenome assembly and assigned initially as baculovirus-like sequences [20]. The inner circle represents the GC plot (above average: gold; below average: purple). The grey triangles flanked by arrows indicate unresolved sequences.

## 2.5. Nucleotide Sequence Accession Number

The nucleotide sequence of the AmFV genome was deposited in GenBank under accession KR819915.



### 3. Results and Discussion

#### 3.1. Assembly and Nucleotide Sequence Analysis of AmFV

The sequencing output was 25.67 million read pairs with a median size of 404 bp. Trimming and assembly resulted in the selection of an optimal 93-bp kmer assembly, based on the highest N50 (155,624 bp) and proportion of reads mapping to contigs (90.1%). The resulting 276 contigs had a total length of 3.9 Mbp. Scaffolding by Velvet produced two major scaffolds of 367,058 and 131,360 nucleotides. These scaffolds could furthermore be joined by PCR, such that the junctions between the scaffolds could be sequenced from both directions by Sanger sequencing of the PCR fragments. However, the internal regions of the gaps between the scaffolds, of approximately 250 and 1200 nucleotides, could not be resolved due to their high GC content (>70%) and the presence of repeated sequences. These results led us to infer that the AmFV genome consists of a circular dsDNA molecule of approximately 498,500 nucleotides, with an average GC content of 50.8% (Figure 2). The AmFV genome therefore appears intermediate in size between mimiviruses (>1 Mb) [26] and the White Spot Syndrome Virus (WSSV) [27] or the largest entomopoxviruses [28,29].

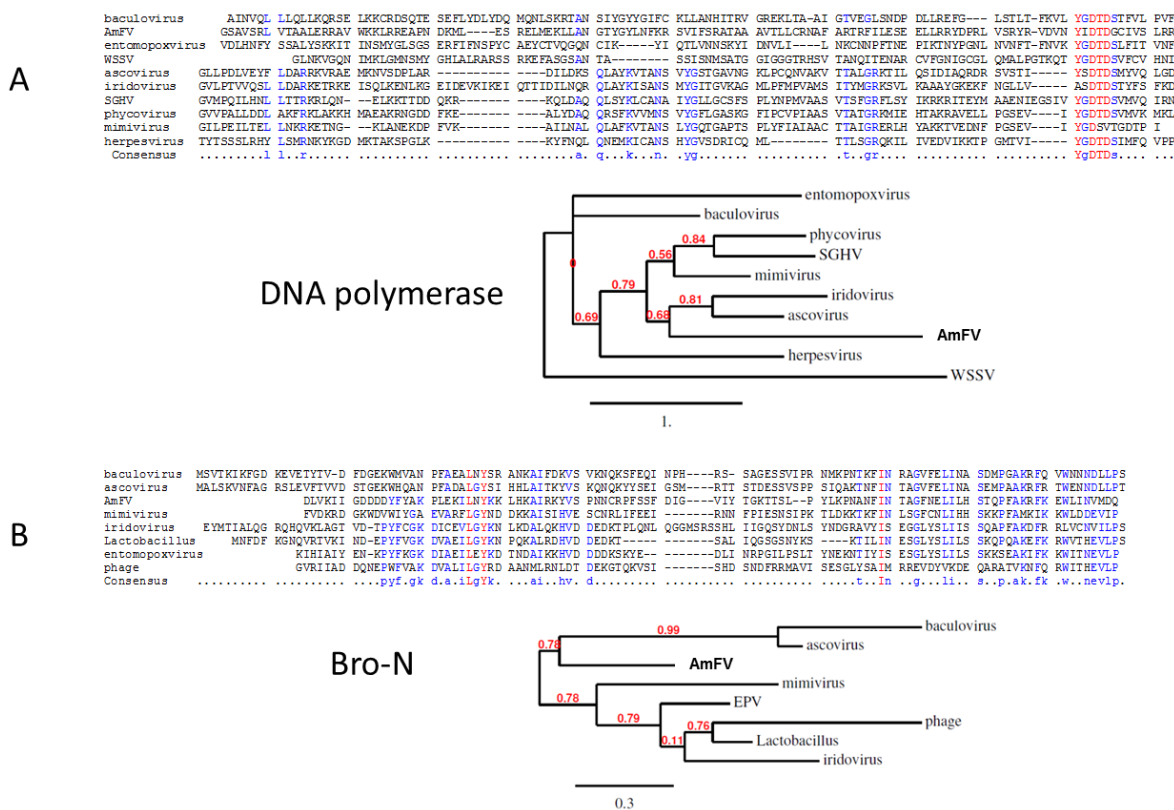
A plot of kmer abundance is shown in Supplemental File 2a. Two distinct modes of coverage are evident around 160× and 380×. Based on the inflection points of the curve in Figure 2a, kmers in the coverage range of 50–220 and 250–750 were investigated separately as likely deriving from two distinct genomes present in the isolate. The lower coverage mode included approximately four times as many distinct kmers (5,905,345 vs. 1,547,736), with a GC content of 36.7% vs. 61.7%. Each batch of kmers was assembled separately and aligned with BLASTX to the UniRef90 database (matches summarized in Supplemental File 2b). The results indicate that the first kmer coverage mode represents genomic sequence of an *Arsenophonus* bacterium, because a large majority of best matches were to *Arsenophonus nasoniae*, generally with protein similarities >90%. The second, high-GC kmer population appears to represent a plasmid and/or phage associated with *Arsenophonus*, because *Arsenophonus* is again the most frequently matched taxon, but with a greater representation of other taxa and the spectrum of proteins identified is narrower and dominated by those associated with conjugation, replication, and phage assembly. Similar BLASTX matches as those found for the AmFV genome model were also found in this second pool of kmers, although a minority of matches overall. Thus, the isolation and assembly of AmFV appears to have been complicated by co-occurrence of these other genomes, explaining the need for substantial laboratory confirmation and scaffolding of the AmFV genome model despite high total coverage of paired-end reads. We are nevertheless confident about the accuracy of the assembled AmFV sequence due to the considerable overlap with US contigs isolated from the *V. destructor* metagenome (Figure 2).

#### 3.2. Gene Content Analysis

A total of 247 putative non overlapping coding DNA sequences with a minimum length of 100 amino acids were assumed to encode putative proteins. If these 247 represent the total coding DNA of the genome, then the coding density of the AmFV genome is about 65%. The coding density value of AmFV is then intermediate regarding values encountered in other large dsDNA viruses, which range for instance between 90.5% (Mimivirus) and 26.9% (Bracovirus) [30]. The predicted ORFs are evenly

distributed on both strands and only 40 ORFs (16%) had threshold similarity to proteins in the public sequences databases (see supplementary file S3). Of these, 28 ORFs showed significant homology with proteins of large DNA viruses, 13 of which with baculoviruses [31] (Table 1 and Table S3).

The most conserved ORFs in terms of percent amino-acid identity are those associated with nucleotide metabolism, e.g., AmFV\_116 and AmFV\_221, encoding putative subunits of a ribonucleotide reductase. This enzyme is highly conserved among living organisms and catalyzes the formation of deoxyribonucleotides from ribonucleotides for the synthesis of DNA. Other highly conserved genes are those involved in DNA synthesis, such as AmFV\_28 encoding a putative thymidylate synthase, and AmFV\_76, encoding a putative 1954 amino acid DNA polymerase characterized by a type B DNA polymerase (POLBc) domain typical of large dsDNA virus families (Figure 3A).



**Figure 3.** Multiple sequence alignments and phylogenetic trees for the DNA polymerase (A) and Bro-N (B) domains. Branch support values are indicated in red. GenBank entries from top to bottom: A, gil365199561l, gil1063688l, gil409978642l, gil589287870l, gil9964364l, gil15021419l, gil187903038l, gil312233904l, gil2702251l, and gil311977705l; and B, gil325152622l, gil21668326l, gil55416627l, gil9631452l, gil672972322l, gil736994060l, and gil660515707l.

**Table 1.** Putative homologies of methionine initiated ORFs, with a minimum length of 100 amino acids, identified throughout the AmFV genome (best match against Pfam-A database with Hmmer3 (**left**) and corresponding best BLASTP match against viral sequences (**right**)). Records with e-value under threshold are indicated with a dashed line. The percent identity at the amino-acid level between Swiss and US protein sequences are noted on the right.

Putative function	Best Pfam-A database match with HMMer3 (0.1 cut off)					Best match with viral sequences database (BLASTP search taxid 10239, 1.0e-5 cut off)					AmFV (USA) % identity
	AmFV ORF	Size (aa)	Pfam code domain	e-value	Pfam n <sup>o</sup>	CDS	e-value	Length / %similarity	Score	Accession number	
DNA replication and nucleotide metabolism	AmFV_76	1954	DNA_pol_B	6.5e-17	PF00136.16	-	-	-	-	-	98.1
	AmFV_97	1602	DNA ligase	2.8e-08	PF04675	Hyp. protein [ <i>P. bursaria</i> Chlorella virus AR158]	4.0e-12	509 / 53% (61/115)	72	YP_001498814.1	93.7
	AmFV_116	878	Ribonuc_red_IgC	5.9e-180	PF02867.10	RR1 [ <i>S. litura</i> nucleopolyhedrovirus]	0.0	770 / 42% (521/883)	711	NP_258291.1	99.7
	AmFV_221	2332	Ribonuc_red_sm	3.4e-96	PF00268.16	RR2B [ <i>S. litura</i> nucleopolyhedrovirus]	7.e-56	333 / 72% (132/182)	203	NP_258331.1	97.8
Virion structure and morphogenesis	AmFV_28	610	Thymidylat_synt	2.0e-86	PF00303.14	wsv067 [Shrimp white spot syndrome virus]	3.0e-82	289 / 63% (188/294)	265	NP_477589.1	98.4
	AmFV_125	506	Capsid_NCLDV	1.5e-02	PF04451.7	Hyp. protein [Organic Lake phycodnavirus]	5.0e-10	238 / 40% (116/283)	65	ADX05786.1	99.4
	AmFV_102	401	Baculo_44	1.3e-16	PF04631.7	PIF-2 [ <i>G. bimaculatus</i> nudivirus]	2.0-15	378 / 48% (84/173)	82	YP_001111333.1	99.8
	AmFV_90	279	Pif-3	4.0e-06	PF05006.7	PIF3 [ <i>E. ello</i> granulovirus]	9.0e-06	188 / 39% (64/164)	50	YP_009091870.1	99.6
	AmFV_61	1057	Pif-1	1.6e-17	PF05092.7	PIF1 [ <i>S. littoralis</i> nucleopolyhedrovirus]	3.0-13	525 / 43% (99/230)	77	AGE89974.1	95.7
	AmFV_62	829	Pif-1	5.8e-25	PF05092.7	DekiORF31 [ <i>D. kikuchii</i> nucleopolyhedrovirus]	9.0e-19	536 / 48% (109/226)	95	AFS51909.1	98.6
	AmFV_79	1196	Baculo_p74	6.6e-10	PF08404.5	P74 [ <i>B. mori</i> nucleopolyhedrovirus]	1.0e-9	645 / 50% (72/144)	68	NP_047536.1	96.8
Unknown	AmFV_110	626	Bro N	1.82e-05	PF02498	BRO-C [ <i>M. configurata</i> nucleopolyhedrovirus]	8.0e-11	326 / 50% (110/219)	69	NP_689249.1	96.9
	AmFV_112	498	Bro N	1.5e-13	PF02498.12	BRO-B [ <i>C. chalcites</i> nucleopolyhedrovirus]	7.0e-17	628 / 48% (134/279)	85	YP_249673.1	97.8
	AmFV_108	662	Bro N	1.7e-09	PF02498.12	BRO-M [ <i>L. xyli</i> nucleopolyhedrovirus]	2.0e-12	474 / 46% (95/204)	75	YP_003517887.1	99.2
	AmFV_17	1313	Bro N	6.5e-09	PF02498.12	DekiORF51 [ <i>D. kikuchii</i> nucleopolyhedrovirus]	8.0e-08	480 / 49% (55/112)	62	AFS51929.1	90.9
	AmFV_9	181	ns	ns	ns	BRO-D [ <i>C. chalcites</i> nucleopolyhedrovirus]	7.0e-07	429 / 56% (42/75)	53	YP_249718.1	100
	AmFV_77	432	ns	ns	ns	BRO-6 [ <i>S. litura</i> granulovirus]	1.0e-05	485 / 42% (58/136)	52	YP_001257066.1	98.5
	AmFV_197	898	Chitin binding 3	6.8e-25	PF03067.10	-	-	-	-	-	96.1
	AmFV_104	596	Collagen	2.2e-07	PF01391.13	collagen repeat [Bacillus phage phBC6A52]	9.0e-07	536 / 60% (46/76)	57	NP_852574.1	94.0
	AmFV_69	422	ns	ns	ns	collagen-like protein [ <i>A. polyphaga</i> mimivirus]	3.0e-07	1392 / 54% (54/99)	58	YP_003987190.1	95.9
	AmFV_36	501	Pacifastin I	1.6e-10	PF05375.8	-	-	-	-	-	94.3
	AmFV_3	963	Abhydrolase_3	1.3e-09	PF07859.8	-	-	-	-	-	96.7
	AmFV_6	1354	Peptidase_M10	1.2e-10	PF00413.19	-	-	-	-	-	92.2
	AmFV_223	627	Peptidase_M10	1.3e-10	PF00413.19	-	-	-	-	-	97.9
	AmFV_12	576	Kinesin	1.1e-67	PF00225.18	-	-	-	-	-	99.8



Also identified in the AmFV genome were several putative orthologs of the baculovirus *pif-1*, *pif-2*, *pif-3* and *p74* genes. These genes belong to baculovirus core gene set and are required for oral infectivity [31]. They are also present in other DNA viruses infecting arthropods, such as nudiviruses [32,33], salivary glands hypertrophy viruses (SGHVs) [34,35], WSSV [27] and bracoviruses [36–38]. It is not known whether these gene families are required for infectivity in these other viral groups, nor whether they were acquired horizontally or instead reflect common ancestry [35]. *Pif* genes are absent from entomopoxvirus genomes, however, even though these viruses are also orally transmitted [39]. Other elements typical of baculoviruses that were also found in the AmFV genome are the *Bro* genes, of which at least six were found in the AmFV genome. BRO proteins, whose functions are still unknown, contain a DNA binding domain and belong to a multigene family that is wide spread among large insect viruses, including entomopoxviruses and iridoviruses, as well as in dsDNA phage and prokaryotic class II transposons [40,41].

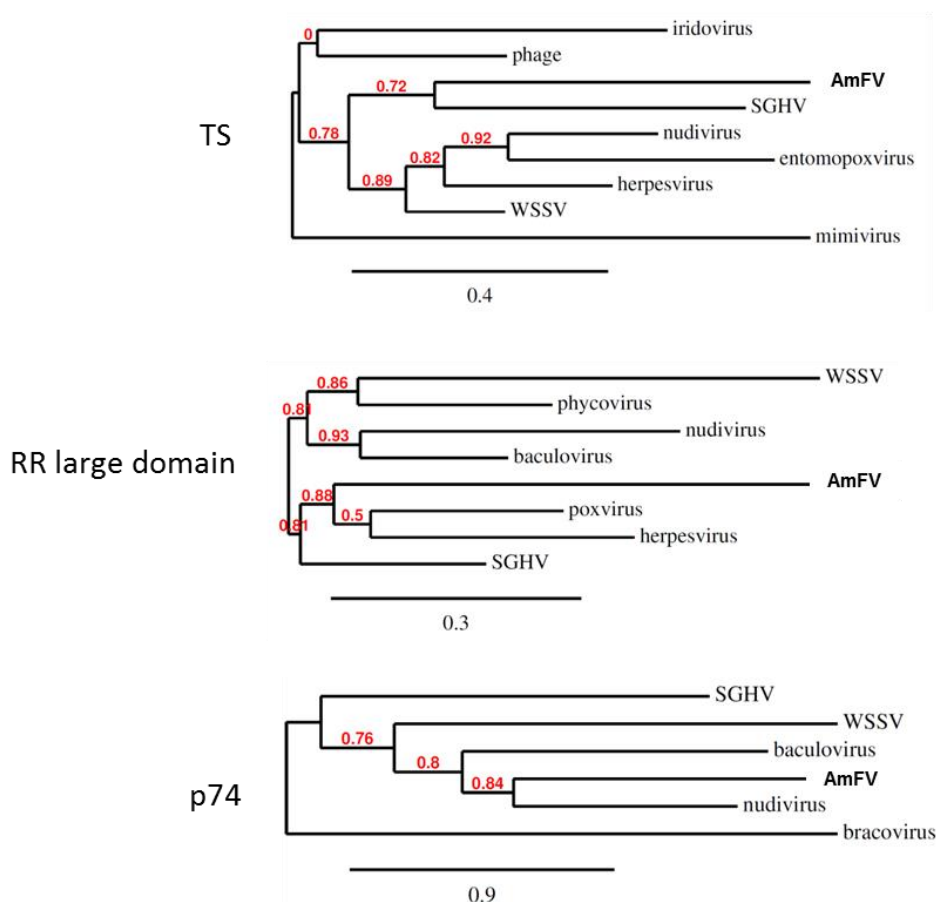
Several additional potential virulence factors were identified in the AmFV genome. Such factors encompass a range of functions that affect infectivity and host-virus interactions and often include degradative enzymes (chitinases, proteases) that are common in many insect virus genomes [42]. The AmFV\_197 encodes a chitin-binding domain (although it lacks a domain for the corresponding chitinase enzyme) that may play a role during the early infectious process. Chitin-binding domains are also found in the fusolin proteins of entomopoxviruses and baculoviruses that facilitate the passage of the virus through the peritrophic membrane in the gut [43,44] and similar genes are also present in bacterial pathogens of honeybees [45]. The AmFV\_36 protein contains four pacifastin motifs, as well as a transmembrane domain. Pacifastin belongs to a family of serine proteinase inhibitors that is found in arthropods but so far not in viruses. It has recently been shown that pacifastin regulates the prophenoloxidase cascade in crabs [46]. This class of serine inhibitors also seems to be involved in regulating certain physiological processes in insects [47]. If the AmFV pacifastin has a similar function, it may interfere with the immune system of its host *A. mellifera*. AmFV\_6 and AmFV\_223 are proteins with zinc-dependent metalloprotease domains. AmFV\_6 is a putative ortholog of a metalloprotease found in *Pieris rapae* granulovirus. Metalloproteases feature often as virulence factors for many pathogens, such as the bacterium *Paenibacillus larvae*, the causative agent of American foulbrood disease in honeybees [48]. AmFV\_12 has homology with a domain present in the drosophila kinesin gene (e-value:  $2.e^{-42}$ , see Supplementary File S3). Kinesin belongs to a class of motor proteins that has been shown to facilitate intracellular trafficking of viral components along the microtubule cytoskeleton [49,50]. The presence of a putative homologue of an insect cellular gene in the AmFV genome suggests that AmFV may have evolved a mutualistic association with its host, involving horizontal transfer of genes between host and virus, as it is known to occur with certain other large DNA viruses, such as Bracoviruses [37,51], Baculoviruses and Poxviruses [41]. This possibility requires more detailed investigations. There are also several ORFs that contain collagen-like domains. Such domains are found in a variety of large nucleocytoplasmic DNA viruses and may be associated with virulence. It has been shown that collagen repeated sequences are a key determinant of herpes transforming activity [52] and that collagen is expressed at the surface of the mimivirus, a large DNA virus infecting amoebae, and that this can induce an immune response in mice [53].

### 3.3. Sequence Variation between European and North American AmFV Isolates

The baculovirus-like contigs identified in *V. destructor* by Cornman *et al.*, 2010 [20] showed very high nucleotide identity with the AmFV genome sequence from Switzerland, as illustrated in Figure S1. A majority of the ORFs were also identified in this set of varroa-derived contigs from USA, most of them with very high identity scores (Table 1). For example, the putative DNA polymerase was highly conserved between the two AmFV isolates, with just 17 substitutions and six insertions/deletions along the 1954 amino acid sequence, in addition to differences in the number of repetitions of GN/S residues that may represent a hinge region between different domains of the polymerase. These data strongly suggest that the sequences from the USA and Switzerland belong to the same virus, AmFV. One significant difference concerns the putative reverse transcriptase identified by Cornman *et al.*, 2010, indicative of the presence of a retrotransposon, which was absent in the AmFV genome from Switzerland. Mobile elements are rare in large DNA viruses and have so far only been reported in bracoviruses belonging to the *Cotesia* and *Glyptapanteles* genera [54,55] and the *Melanoplus sanguinipes* entomopoxvirus [56].

### 3.4. AmFV Classification

The analysis of the AmFV genome sequence highlighted a number of features that make it difficult to assign it to a particular group of viruses. The presence of putative orthologs of baculoviruses suggests a relationship between AmFV and *Baculoviridae* (Table 1). However, this overlap is restricted to a few genes and the large size of the genome and the presence of a high proportion of non-coding sequences are inconsistent with a baculovirus designation. In addition, many other coding sequences present in the AmFV genome, such as the genes involved in DNA replication, display greater similarities with their homologues in other large DNA viruses such as *Poxviridae*, *Herpesviridae* or *Phycoviridae* (Figures 3 and 4). For instance, the phylogenetic analyses of the BroN domain of the AmFV\_112 ORF (Figure 3B) and the p74 protein (Figure 4) positioned AmFV close to the baculoviruses, while the DNA polymerase domain of AmFV\_76 placed it within the ascovirus and iridovirus cluster (Figure 3A). Likewise, the ribonucleotide reductase IgC domain of AmFV\_116 ORF clustered with poxviruses and herpesviruses (Figure 4). This pattern is consistent with networks of multiple horizontal gene transfers (HGT) between viruses infecting the same host, or gene acquisition from the host or its microbiota [41,57]. Indeed, many genes in the AmFV genome have putative viral or bacterial homologues, such as the Bro genes, while others are homologues of insect genes, such as the AmFV\_12 ORF which has a kinesin domain also present in *Drosophila pseudoobscura*, and the AmFV\_36 ORF which has homologies with the precursor of pacifastin venom protein 1 of *Nasonia vitripennis* (Table S3, nr database). The dispersion of core genes throughout the AmFV genome also support potential HGT processes, as does the presence of a retrotransposon in the US AmFV isolate, since such mobile elements are known to be frequently involved in the assimilation of exogenous genes [57–59]. The variable GC content observed throughout the AmFV genome (Figure 2) is also consistent with putative HGT processes that may have been responsible for the still unclear taxonomic position of the AmFV genome.

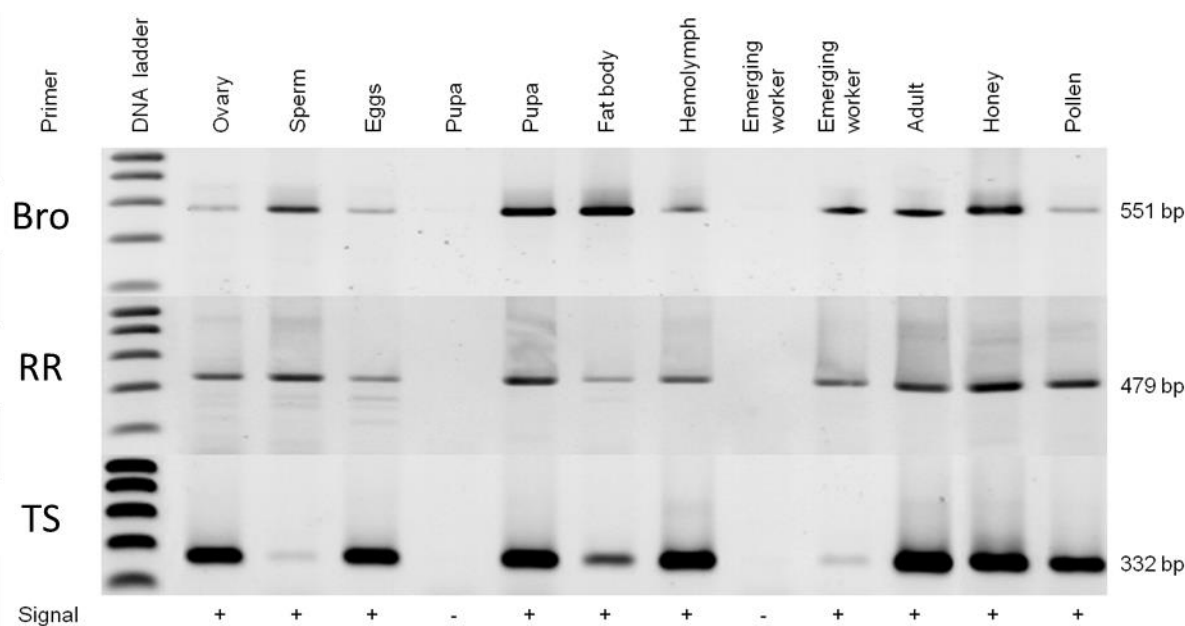


**Figure 4.** Phylogenetic analyses. Three were built from thymidylate synthase (TS), ribonucleotide reductase (RR) and p74 proteins. Branch support values are indicated in red. GenBank entries from top to bottom: TS, gil589287845l, gil485725392l, gil187903049l, gil370702981l, gil9631429l, gil551484996l, gil417072295l, and gil55417112l; RR large domain, gil15021484l, gil347481982l, gil370702993l, gil294471329l, gil22164662l, gil422933669l, gil187903102l; p74, gil187903076l, gil17016513l, gil9630072l, gil370703052l, and gil531034105l.

### 3.5. AmFV Distribution in Honeybee Colonies

A PCR-based diagnostic assay was developed from three coding sequences with high nucleotide identity between the AmFV isolates from Switzerland and USA: (a) the putative thymidylate synthase gene (AmFV\_28), (b) one of the Bro genes (AmFV\_112) and (c) the ribonucleotide reductase gene (AmFV\_221). The alignments between the Swiss and US AmFV isolates for these sequences are provided in Figure S1. Using these PCR assays, pools of bees from different colonies were analyzed to compare the prevalence of AmFV in the USA and Switzerland. AmFV was detected in 64% of the Swiss colonies tested (N = 25) and in 100% of the US colonies (N = 15). These results are consistent with classical reports of a relatively high prevalence and wide geographic distribution of AmFV [6,60]. A second screening was conducted to determine the distribution of AmFV in different honeybee organs and bee products found in the colony (Figure 5). The detection of AmFV in pupae, emerging workers and adults suggests that AmFV can infect both the brood and adult stages. Moreover, AmFV infections seem

to remain through metamorphosis, thereby suggesting that the virus is not restricted to the alimentary tract. This is supported by the detection of AmFV in the hemolymph of adult workers. AmFV was furthermore also detected in honey and pollen, indicating a possible oral-fecal transmission route. The queen ovaries, eggs and sperm from the drone endophalli were also positive for AmFV, which is indicating considerable potential for either vertical (transovarial) or venereal transmission of this virus, similar to deformed wing virus [61]. This is also supported by the high prevalence of AmFV in queens (11/13 mated queens analyzed were found positive). Likewise, AmFV seems to be present in winter bees, as 12/19 individual overwintering bees from California and Florida were found positive. There was no difference in AmFV prevalence between colonies affected or not affected by Colony Collapse Disorder (CCD), with a combined prevalence of 84% (N = 75). AmFV could only be detected in 3/43 individual *V. destructor* mites and 3/29 pooled mite samples (100 mites/sample), concurring with prior assessments [20]. Moreover, AmFV was also highly prevalent in *V. destructor*-free colonies from the island of Ouessant, France (N = 15), indicating that AmFV presence and prevalence are not closely related to *V. destructor* mite infestation of honeybee colonies. These data suggest that *V. destructor* is probably not a major factor in AmFV prevalence and transmission. The presence of AmFV in *V. destructor* mite samples, even with high abundance, can easily be explained by passive acquisition, since AmFV sheds large amounts of particles and nucleoproteins into the bee hemolymph (“milky hemolymph” symptom), on which the ectoparasite mite feeds. The frequent detection of AmFV in bee guts and food products suggests that oral transmission may be important, consistent with classical observations [7] while the high frequency of AmFV in mated queens, ovaries and sperm suggests a possible vertical transmission route, although previous studies failed to show vertical transmission from artificially AmFV-infected queens [2].



**Figure 5.** Agarose gel electrophoresis showing AmFV PCR amplicons produced from different individuals (larvae, pupae or adults bees), queen bee tissues, drone sperm or bee products, using three different primer pairs (the thymidylate synthase (TS) primers were described by Cornman *et al.* (2010) [20]).

#### 4. General Conclusions

We have presented a model of the AmFV genome based on deep sequencing of an enriched extract of semi-pure filamentous virus, identified by classical symptoms, supplemented by Sanger sequencing of specific fragments bridging different assembly scaffolds. Paired-end reads strongly support the accuracy of the final genome model, although they were not able to bridge the long, GC-rich repetitive region that was ultimately scaffolded by PCR. Despite its large size, the presence of a large proportion of novel protein sequence, and its relatively low gene density, AmFV is well conserved at the nucleotide level between two isolates from different continents. This conservation may reflect the high fidelity of the AmFV DNA polymerases and/or the facilitated dispersal of AmFV by the global bee trade. The current genome model can be used as a foundation for functional studies, assays development, quantitative assessments with qPCR, and comparative genomics through metagenomic sequencing of other bee samples.

#### Supplementary Information

**Figure S1.** Alignment of three AmFV genes used for designing PCR primers. AmFV-US sequences were identified in the USA by Cornman *et al.*, 2010 [20]. Primer locations are underlined.

**Figure S2.** Details about NGS sequencing data.

**Figure S2a.** The number of unique 59-bp kmers occurring in the filtered reads (vertical axis) is plotted as a function of the number of times that they are present (horizontal axis). As different genomes in a heterogeneous mixture are often present in distinctly different copy numbers, these kmer plots and subsequent analysis of the properties of kmers at different coverage levels allows a partial disentangling of those genomes. Stochasticity, sequence similarity, polymorphism, extraction/sequencing bias, and repetitive components of genomes also influence the abundance of kmers and thus not all kmers are expected to occur proportional to the number of copies of the genome from which they derive.

**Figure S2b.** In total, 59-bp kmers in the coverage range of 50–220 and 250–750 were identified using the program jellyfish and assembled with CLC Genomics v 7.1 at kmer = 59. Best BLASTX matches of the resulting contigs at an e-value threshold of  $e^{-10}$  are shown.

**Table S3.** Blastp hits against (a) the non-redundant protein sequences database (<http://blast.ncbi.nlm.nih.gov/>); (b) same as (a) but restricted to viruses (taxid 10239); and (c) to the UniRef90 database [13]. A list of the AmFV ORFs with corresponding coding sequences is also provided in this file.

**Figure S4.** Multiple amino-acid sequence alignments of thymidylate synthase (TS), ribonucleotide reductase large fragment (RR1), and p74, using the Multalin software [16].

#### Acknowledgments

We are very grateful to Max Bergoin for help in microscopy and for fruitful discussions. We also like to thank two anonymous reviewers for constructive comments. The AmFV sequencing project was financially supported by Agroscope and the Vinetum Foundation.

## Author Contributions

L.G., S.C. and F.C. performed the sequencing analyses. U.H. and J.D.E. were involved in the diagnostic assays. P.N. procured funding via the Vinetum Foundation. L.G., S.C., J.R.M. and P.N. wrote the manuscript.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Wille, H. Septikämien und Mischinfektionen. *Schweiz. Bienenztg.* **1962**, *85*, 222–226.
2. Clark, T.B. A filamentous virus of the honey bee. *J. Invertebr. Pathol.* **1978**, *32*, 332–340. [[CrossRef](#)]
3. Sitaropoulou, N.; Neophytou, E.P.; Thomopoulos, G.N. Structure of the Nucleocapsid of A Filamentous Virus of the Honey Bee (*Apis-mellifera*). *J. Invertebr. Pathol.* **1989**, *53*, 354–357. [[CrossRef](#)]
4. Bailey, L.; Carpenter, J.M.; Woods, R.D. Properties of filamentous virus of the honey bee. *Virology* **1981**, *114*, 1–7. [[CrossRef](#)]
5. Federici, B.A.; Bideshi, D.K.; Tan, Y.; Spears, T.; Bigot, Y. Ascoviruses: Superb Manipulators of Apoptosis for Viral Replication and Transmission. In *Lesser Known Large DsDNA Viruses*; Current Topics in Microbiology and Immunology. Van Etten, J.L., Ed.; Springer-Verlag: Berlin, Germany; Heidelberg, Germany, 2009; Volume 328, pp. 171–196.
6. Allen, M.F.; Ball, B.V. The incidence and world distribution of honey bee viruses. *Bee World* **1996**, *77*, 141–162. [[CrossRef](#)]
7. Bailey, L.; Ball, B.V. *Honey Bee Pathology*; Academic Press: London, UK, 1991.
8. Varis, A.L.; Ball, B.V.; Allen, M. The Incidence of Pathogens in Honey-Bee (*Apis mellifera* L.) Colonies in Finland and Great-Britain. *Apidologie* **1992**, *23*, 133–137. [[CrossRef](#)]
9. Bailey, L.; Ball, B.V.; Perry, J.N. Association of viruses with two protozoal pathogens of the honey bee. *Ann. Appl. Biol.* **1983**, *103*, 13–20. [[CrossRef](#)]
10. Varaldi, J.; Ravallec, M.; Labrosse, C.; Lopez-Ferber, M.; Bouletreau, M.; Fleury, F. Artificial transfer and morphological description of virus particles associated with superparasitism behaviour in a parasitoid wasp. *J. Insect Physiol.* **2006**, *52*, 1202–1212. [[CrossRef](#)] [[PubMed](#)]
11. Zerbino, D.R.; Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **2008**, *18*, 821–829. [[CrossRef](#)] [[PubMed](#)]
12. Marçais, G.; Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **2011**, *27*, 764–770. [[CrossRef](#)] [[PubMed](#)]
13. Suzek, B.E.; Wang, Y.; Huang, H.; McGarvey, P.B.; Wu, C.H.; UniProt Consortium. UniRef clusters: A comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **2015**, *31*, 926–932. [[CrossRef](#)] [[PubMed](#)]
14. Rice, P.; Longden, I.; Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genetics* **2000**, *16*, 276–277. [[CrossRef](#)]



15. Bocs, S.; Cruveiller, S.; Vallenet, D.; Nuel, G.; Médigue, C. AMIGENE: Annotation of Microbial Genes. *Nucleic Acids Res.* **2003**, *13*, 3723–3726. [[CrossRef](#)]
16. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881–10890. [[CrossRef](#)] [[PubMed](#)]
17. Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.F.; Guindon, S.; Lefort, V.; Lescot, M.; *et al.* Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **2008**, *36*, W465–W469. [[CrossRef](#)] [[PubMed](#)]
18. Schultz, J.; Milpetz, F.; Bork, P.; Ponting, C.P. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5857–5864. [[CrossRef](#)] [[PubMed](#)]
19. Carver, T.; Thomson, N.; Bleasby, A.; Berriman, M.; Parkhill, J. DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics (Oxford, England)* **2009**, *25*, 119–120. [[CrossRef](#)] [[PubMed](#)]
20. Cornman, R.S.; Schatz, M.C.; Spencer, J.; Chen, Y.P.; Pettis, J.; Hunt, G.; Bourgeois, L.; Elsik, C.; Anderson, D.; Grozinger, C.M.; *et al.* Genomic survey of the ectoparasitic mite *Varroa destructor*, a major pest of the honey bee *Apis mellifera*. *BMC Genomics* **2010**, *11*, e602. [[CrossRef](#)] [[PubMed](#)]
21. Chevreux, B.; Pfisterer, T.; Drescher, B.; Driesel, A.J.; Müller, W.E.G.; Wetter, T.; Suhail, S. Using the miraEST Assembler for Reliable and Automated mRNA Transcript Assembly and SNP Detection in Sequenced ESTs. *Genome Res.* **2004**, *14*, 1147–1159. [[CrossRef](#)] [[PubMed](#)]
22. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)] [[PubMed](#)]
23. Picard, A Set of Tools (in Java) for Working with Next Generation Sequencing Data in the BAM (<http://samtools.sourceforge.net>) Format. Available online: <http://broadinstitute.github.io/picard/> (accessed on 7 July 2015).
24. Carreck, N.L.; Andree, M.; Brent, C.S.; Cox-Foster, D.; Dade, H.A.; Ellis, J.D.; Hatjina, F.; van Engelsdorp, D. Standard methods for *Apis mellifera* anatomy and dissection. *J. Apic. Res.* **2013**, *52*. [[CrossRef](#)]
25. Gauthier, L.; Ravallec, M.; Tournaire, M.; Cousserans, F.; Bergoin, M.; Dainat, B.; de Miranda, J.R. Viruses associated with ovarian degeneration in *Apis mellifera* L. queens. *PLoS ONE* **2011**, *6*, e16217. [[CrossRef](#)] [[PubMed](#)]
26. Claverie, J.M.; Abergel, C.; Ogata, H. Mimivirus. *Curr. Top. Microbiol. Immunol.* **2009**, *328*, 89–121. [[PubMed](#)]
27. Van Hulten, M.C.W.; Witteveldt, J.; Peters, S.; Kloosterboer, N.; Tarchini, R.; Fiers, M.; Sandbrink, H.; Klein Lankhorst, R.; Vlak, J.M. The white spot syndrome virus DNA genome sequence. *Virology* **2001**, *286*, 7–22. [[CrossRef](#)] [[PubMed](#)]
28. Bawden, A.L.; Glassberg, K.J.; Diggans, J.; Shaw, R.; Farmerie, W.; Moyer, R.W. Complete Genomic Sequence of the *Amsacta moorei* Entomopoxvirus: Analysis and Comparison with Other Poxviruses. *Virology* **2000**, *274*, 120–139. [[CrossRef](#)] [[PubMed](#)]
29. Thézé, J.; Takatsuka, J.; Li, Z.; Gallais, J.; Doucet, D.; Arif, B.; Nakai, M.; Herniou, E.A. New insights into the evolution of *Entomopoxvirinae* from the complete genome sequences

- of four entomopoxviruses infecting *Adoxophyes honmai*, *Choristoneura biennis*, *Choristoneura rosaceana*, and *Mythimna separata*. *J. Virol.* **2013**, *87*, 7992–8003. [[CrossRef](#)] [[PubMed](#)]
30. Desjardins, C.; Eisen, J.A.; Nene, V. New evolutionary frontiers from unusual virus genomes. *Genome Biol.* **2005**, *6*, e212. [[CrossRef](#)] [[PubMed](#)]
31. Herniou, E.A.; Arif, B.M.; Becnel, J.J.; Blissard, G.W.; Bonning, B.; Harrison, R.; Jehle, J.A.; Theilmann, D.A.; Vlak, J.M. Baculoviridae. In *Virus Taxonomy*; King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., Eds.; Elsevier: Oxford, UK, 2011; pp. 163–174.
32. Wang, Y.; Jehle, J.A. Nudiviruses and other large, double-stranded circular DNA viruses of invertebrates: New insights on an old topic. *J. Invertebr. Pathol.* **2009**, *101*, 187–193. [[CrossRef](#)] [[PubMed](#)]
33. Jehle, J.A. Nudiviruses: Their biology and genetics. In *Insect Virology*; Asgari, S., Johnson, K.N., Eds.; School of Biological Sciences, The University of Queensland: St Lucia QLD, Australia, 2010; pp. 153–170, 436.
34. Lietze, V.U.; Abd-Alla, A.M.; Vreysen, M.J.; Geden, C.J.; Boucias, D.G. Salivary gland hypertrophy viruses: A novel group of insect pathogenic viruses. *Annu. Rev. Entomol.* **2011**, *56*, 63–80. [[CrossRef](#)] [[PubMed](#)]
35. Jehle, J.A.; Abd-Alla, A.M.; Wang, Y. Phylogeny and evolution of Hytrosaviridae. *J. Invertebr. Pathol.* **2013**, *112*, S62–S67. [[CrossRef](#)] [[PubMed](#)]
36. Espagne, E.; Dupuy, C.; Huguet, E.; Cattolico, L.; Provost, B.; Martins, N.; Poirie, M.; Periquet, G.; Drezen, J.M. Genome sequence of a polydnavirus: Insights into symbiotic virus evolution. *Science* **2004**, *306*, 286–289. [[CrossRef](#)] [[PubMed](#)]
37. Bézier, A.; Herbinière, J.; Lanzrein, B.; Drezen, J.M. Polydnavirus hidden face: The genes producing virus particles of parasitic wasps. *J. Invertebr. Pathol.* **2009**, *101*, 194–203. [[CrossRef](#)] [[PubMed](#)]
38. Wetterwald, C.; Roth, T.; Kaeslin, M.; Annaheim, M.; Wespi, G.; Heller, M.; Mäser, P.; Roditi, I.; Pfister-Wilhelm, R.; Bézier, A.; *et al.* Identification of bracovirus particle proteins and analysis of their transcript levels at the stage of virion formation. *J. Gen. Virol.* **2010**, *91*, 2610–2619. [[CrossRef](#)] [[PubMed](#)]
39. Perera, S.; Zhen, L.; Pavlik, L.; Arif, B. Entomopoxviruses. In *Insect Virology*; Asgari, S., Johnson, K.N., Eds.; School of Biological Sciences, The University of Queensland: St Lucia QLD, Australia, 2010; pp. 83–102, 436.
40. Bideshi, D.K.; Renault, S.; Stasiak, K.; Federici, B.A.; Bigot, Y. Phylogenetic analysis and possible function of bro-like genes, a multigene family widespread among large double-stranded DNA viruses of invertebrates and bacteria. *J. Gen. Virol.* **2003**, *84*, 2531–2544. [[CrossRef](#)] [[PubMed](#)]
41. Thézé, J.; Takatsuka, J.; Nakai, M.; Arif, B.; Herniou, E.A. Gene acquisition convergence between entomopoxviruses and baculoviruses. *Viruses* **2015**, *13*, 1960–1974. [[CrossRef](#)] [[PubMed](#)]
42. Ishimwe, E.; Hodgson, J.J.; Clem, R.J.; Passarelli, A.L. Reaching the melting point: Degradative enzymes and protease inhibitors involved in baculovirus infection and dissemination. *Virology* **2015**, *479C–480C*, 637–649. [[CrossRef](#)] [[PubMed](#)]

43. Xu, J.; Hukuhara, T. Enhanced infection of a nuclear polyhedrosis virus in larvae of the armyworm, *Pseudaletia separata*, by a factor in the spheroids of an entomopoxvirus. *J. Invertebr. Pathol.* **1992**, *60*, 259–264. [[CrossRef](#)]
44. Mitsuhashi, W.; Kawakita, H.; Murakami, R.; Takemoto, Y.; Saiki, T.; Miyamoto, K.; Wada, S. Spindles of an entomopoxvirus facilitate its infection of the host insect by disrupting the peritrophic membrane. *J. Virol.* **2007**, *81*, 4235–4243. [[CrossRef](#)] [[PubMed](#)]
45. Garcia-Gonzalez, E.; Poppinga, L.; Funfhaus, A.; Hertlein, G.; Hedtke, K.; Jakubowska, A.; Genersch, E. Paenibacillus larvae Chitin-Degrading Protein PICBP49 Is a Key Virulence Factor in American Foulbrood of Honey Bees. *PLoS Pathog.* **2014**, *10*, e7. [[CrossRef](#)] [[PubMed](#)]
46. Liu, Y.; Cui, Z.; Shi, G.; Luo, D.; Wang, S.; Wang, C. PtPLC, a pacifastin-related inhibitor involved in antibacterial defense and prophenoloxidase cascade of the swimming crab *Portunus trituberculatus*. *Fish Shellfish Immunol.* **2015**, *43*, 36–42. [[CrossRef](#)] [[PubMed](#)]
47. Breugelmans, B.; Simonet, G.; van Hoef, V.; van Soest, S.; Vanden Broeck, J. Identification, distribution and molecular evolution of the pacifastin gene family in Metazoa. *BMC Evol. Biol.* **2009**, *9*, e97. [[CrossRef](#)] [[PubMed](#)]
48. Antunez, K.; Arredondo, D.; Anido, M.; Zunino, P. Metalloprotease production by *Paenibacillus larvae* during the infection of honeybee larvae. *Microbiology* **2011**, *157*, 1474–1480. [[CrossRef](#)] [[PubMed](#)]
49. Greber, U.F.; Way, M. A superhighway to virus infection. *Cell* **2006**, *124*, 741–754. [[CrossRef](#)] [[PubMed](#)]
50. Marsh, M.; Helenius, A. Virus entry: Open sesame. *Cell* **2006**, *124*, 729–740. [[CrossRef](#)] [[PubMed](#)]
51. Bézier, A.; Louis, F.; Jancek, S.; Periquet, G.; Thézé, J.; Gyapay, G.; Musset, K.; Lesobre, J.; Lenoble, P.; Dupuy, C.; *et al.* Functional endogenous viral elements in the genome of the parasitoid wasp *Cotesia congregata*: Insights into the evolutionary dynamics of bracoviruses. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **2013**, *368*, e20130047. [[CrossRef](#)] [[PubMed](#)]
52. Choi, J.K.; Ishido, S.; Jung, J.U. The collagen repeat sequence is a determinant of the degree of herpesvirus saimiri STP transforming activity. *J. Virol.* **2000**, *74*, 8102–8110. [[CrossRef](#)] [[PubMed](#)]
53. Shah, N.; Hülsmeier, A.J.; Hochhold, N.; Neidhart, M.; Gay, S.; Hennet, T. Exposure to mimivirus collagen promotes arthritis. *J. Virol.* **2014**, *88*, 838–845. [[CrossRef](#)] [[PubMed](#)]
54. Desjardins, C.A.; Gundersen-Rindal, D.E.; Hostetler, J.B.; Tallon, L.J.; Fadrosh, D.W.; Fuerster, R.W.; Pedroni, M.J.; Haas, B.J.; Schatz, M.C.; Jones, K.M.; *et al.* Comparative genomics of mutualistic viruses of *Glyptapanteles* parasitic wasps. *Genome Biol.* **2008**, *9*, eR183. [[CrossRef](#)] [[PubMed](#)]
55. Drezen, J.M.; Bézier, A.; Lesobre, J.; Huguet, E.; Cattolico, L.; Periquet, G.; Dupuy, C. The few virus-like genes of *Cotesia congregata* bracovirus. *Arch. Insect Biochem. Physiol.* **2006**, *61*, 110–122. [[CrossRef](#)] [[PubMed](#)]
56. Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L. The genome of *Melanoplus sanguinipes* entomopoxvirus. *J. Virol.* **1999**, *73*, 533–552. [[PubMed](#)]

57. Stern, A.; Mayrose, I.; Penn, O.; Shaul, S.; Gophna, U.; Pupko, T. An evolutionary analysis of lateral gene transfer in thymidylate synthase enzymes. *Syst. Biol.* **2010**, *59*, 212–225. [[CrossRef](#)] [[PubMed](#)]
58. Pace, J.K.; Gilbert, C.; Clark, M.S.; Feschotte, C. Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17023–17028. [[CrossRef](#)] [[PubMed](#)]
59. Filée, J. Route of NCLDV evolution: The genomic accordion. *Curr. Opin. Virol.* **2013**, *3*, 595–599. [[CrossRef](#)] [[PubMed](#)]
60. Hartmann, U.; Forsgren, E.; Charrière, J.D.; Neumann, P.; Gauthier, L. Dynamics of *Apis mellifera* Filamentous Virus (AmFV) Infections in Honey Bees and Relationships with Other Parasites. *Viruses* **2015**, *7*, 2654–2667. [[CrossRef](#)] [[PubMed](#)]
61. de Miranda, J.R.; Fries, I. Venereal and vertical transmission of deformed wing virus in honeybees (*Apis mellifera* L.). *J. Invertebr. Pathol.* **2008**, *98*, 184–189. [[CrossRef](#)] [[PubMed](#)]

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).